

АДЛЕР РОБЕРТ,
ПРОФЕССОР ИЗРАИЛЬСКОГО
ТЕХНОЛОГИЧЕСКОГО ИНСТИТУТА (ХАЙФА)

ЭВИНГ ДЖОН,
ИСПОЛНИТЕЛЬНЫЙ ДИРЕКТОР
АМЕРИКАНСКОГО МАТЕМАТИЧЕСКОГО ОБЩЕСТВА В 1995–2008 гг. (США)

ТЕЙЛОР ПИТЕР,
ПРОФЕССОР УНИВЕРСИТЕТА МЕЛЬБУРНА (АВСТРАЛИЯ)

СТАТИСТИКИ ЦИТИРОВАНИЯ*

Доклад Международного математического союза в сотрудничестве с Международным советом промышленной и прикладной математики и Институтом математической статистики

Проанализируем использование данных цитирований и злоупотребление или при оценке научных исследований. Сейчас все распространеннее становится идея о том, что оценка исследовательской деятельности должна осуществляться с помощью «простых и объективных» методов. Эти методы обычно интерпретируются как *библиометрические*, т. е. как данные цитирований, и статистические. Существует убеждение, что статистики цитирований более точны, поскольку оперируют числами, а не сложными суждениями и, следовательно, позволяют обходить субъективность экспертной оценки. Но это убеждение необоснованно.

- Опора на статистические данные не является более точной, если они неправильно используются. Данные статистики могут вводить в заблуждение, если их неправильно понимают. Именно интерпретация и обоснованность статистик цитирований в современной библиометрике основывается на субъективном опыте и интуиции.

- Числа вроде бы «объективны», но эта объективность иллюзорна. Оценка смысла цитирования может быть даже более субъективной, чем экспертная оценка. Поскольку в области цитирований субъективный характер оценки менее очевиден, у тех, кто использует данные цитирований, меньше шансов осознать их ограниченность.

Использование только данных цитирований дает неполное, а зачастую поверхностное понимание научного исследования – понимание, пригодное только тогда, когда оно *подкрепляется другими оценками. Числа отнюдь не лучше, чем разумное суждение.*

Использование данных цитирований для оценки научных исследований означает, что цитаты статистик применяются для ранжирования объектов – журналов, статей, людей, программ и дисциплин часто неправильно.

- Для ранжирования журналов наиболее часто используют импакт-фактор. Это простое среднее, выводимое из распределения цитирований некоторой совокупности статей в журнале, – довольно грубая статистика. Кроме того, при оценке журналов с помощью цитирований обнаруживается много сопутствующих факторов, так что сравнение журналов по импакт-факторам требует большой осторожности. Использовать лишь только импакт-фактор при оценке журнала – все равно что при оценке здоровья человека учитывать только его вес.

- Что касается статей, то вместо того, чтобы при их сравнении полагаться на фактическое количество ссылок, люди часто заменяют его импакт-фактором журналов. Они считают, что более высокий импакт-фактор должен означать более высокий уровень цитирования. Но зачастую это не так! Это широко распространенное неправильное использование статистических данных, с которым следует бороться.

- При сопоставлении отдельных ученых сравнивать полные списки цитирований бывает затруднительно. Как следствие, были предприняты попытки найти простые методы учета, которые выражали бы количественно всю сложность списка цитирований ученого. Наиболее заметным из них является *h-индекс*, который, похоже, становится все более популярным. Но даже эпизодические проверки *h-индекса* и его вариантов показывают, что все они – лишь наивная попытка решить сложную проблему анализа списков цитирований. Они охватывают небольшой объем информации о распределении цитирований ученого, упускают важнейшие данные, оценку научного исследования.

* Игра в цифирь, или Как теперь оценивают труд ученого (сборник статей о библиометрике): пер. с англ. М., 2011. С. 6–38.

Обоснованность таких статистик, как импакт-фактор и *h*-индекс, еще недостаточно изучены. Связь этих методов подсчета с качеством научных исследований иногда устанавливается на основе «опыта». Оправданием для тех, кто полагается на них, является их «доступность». Немногочисленные попытки исследовать эти статистики были направлены исключительно на демонстрацию их корреляции с некоторыми другими мерами качества, а не на определение того, как лучше получить полезную информацию из данных цитирований. Мы не отбрасываем статистики цитирований как инструмент оценки качества научных исследований: данные цитирований и статистики могут нести ценную информацию. Мы допускаем, что оценки должны быть практичными, и по этой причине легко получаемая статистика цитирований почти наверняка станет частью всего процесса. Вместе с тем данные цитирований дают лишь ограниченное и неполное представление о качестве научных исследований, а статистики, полученные на основе данных цитирований, порой неправильно понимаются и используются. Научные исследования слишком важны, чтобы измерять их ценность одним грубым инструментом.

Мы надеемся, что те, кто связан с оценками, познакомятся с комментариями и деталями этого доклада, чтобы не только осознать ограниченность статистик цитирований, но и понять, как их лучше использовать. Если мы устанавливаем высокие стандарты проведения научных исследований, то, безусловно, должны установить столь же высокие стандарты для оценки их качества.

Введение

Научные исследования имеют важное значение. Они определяют прогресс в современном мире и дают надежду на то, что мы можем решить некоторые неразрешимые проблемы, стоящие перед человечеством, – от загрязнения окружающей среды до роста народонаселения. Правительства и учреждения по всему миру осуществляют значительную финансовую поддержку научных исследований. Естественно, они хотят знать, разумно ли инвестируются их деньги; хотят оценивать качество научных исследований, за которые они платят, чтобы принимать обоснованные решения относительно будущих инвестиций.

Это явление далеко не ново: люди оценивали научные исследования и ранее. Ново убеждение, что хорошая оценка должна быть «простой и объективной» и что это может быть достигнуто прежде всего на основе подсчетов (статистики) по данным цитирований, а не с помощью сочетания разных методов, включающих экспертные оценки самих ученых. Эта точка зрения ясно выражена в одном из недавних докладов: «После завершающегося в 2008 г. цикла правительство намерено заменить нынешний метод определения качества университетских научных исследований – UK Research Assessment Exercise (RAE). В центре внимания новой системы будут подсчеты, а не экспертные оценки, и ожидается, что основной индекс качества в этой системе станет библиометрическим (использующим подсчет журнальных статей и их цитирований)».

Те, кто ратуют за объективность в такой простой форме, считают, что научные исследования слишком важны, чтобы полагаться на субъективные суждения, что такие подсчеты вносят ясность в процесс ранжирования. Они полагают, что тщательно подобранные метрики (системы подсчетов) независимы и свободны от систематических ошибок. Но самое главное: они считают, что такие метрики позволяют сравнивать все составляющие научных исследований – журналы, статьи, людей, программы и даже целые дисциплины – просто, эффективно и объективно.

Но такая вера в точность, независимость и эффективность подсчетов неуместна.

- Во-первых, точность этих показателей иллюзорна. Общеизвестно, что статистика может лгать. Злоупотребления в этой сфере распространены и вопиющи. Несмотря на неоднократные попытки что-нибудь исправить в этой сфере, правительства, учреждения да и сами ученые продолжают делать необоснованные или даже ложные заключения, неправильно подсчитав объемы цитирований.

- Во-вторых, использование метода, основанного исключительно на цитированиях, попросту заменяет один вид оценок другим. Вместо экспертной оценки мы получаем субъективную интерпретацию. Те, кто призывает учитывать исключительно подсчеты, основанные на цитированиях, не предполагают, что каждое цитирование означает сомнительный подсчет его «импакт» (влияния).

- В-третьих, хотя статистика дает нам ценную информацию для понимания мира, в котором мы живем, она обеспечивает только его частичное понимание. В современном мире иногда модно провозглашать мистическую веру в то, что численные измерения превосходят другие формы познания. Некоторые ученые пропагандируют статистику цитирований в качестве замены научного исследования, что недопустимо. Мы должны не только использовать статистику правильно – мы должны использовать ее разумно.

Мы спорим не с попыткой оценивать исследование, но, скорее, с требованием, что такие оценки должны опираться на цитирование – «простой и объективный» показатель. Это требование часто интерпретируется как необходимость поиска легких для подсчета цифр, которые выстроят в ряд публикации, или людей, или программы. Исследование же имеет множество целей, как конкретных, так и более отдаленных, и его значимость, естественно, должна оцениваться при помощи не одного, а нескольких критериев. Математики знают, что многие предметы или явления, просто нельзя упорядочить или сравнить. Сравнение часто требует более сложного анализа, который иногда так и не позволяет решить, какой из двух предметов «лучше». Правильный ответ на вопрос «Что лучше?» иногда «Смотря как посмотреть!».

Предложение использовать для оценки качества исследования сразу несколько методов делалось и ранее. Публикации могут быть оценены множеством способов, не только при помощи индекса цитирований. Мерилом профессионального качества исследователя могут быть такие знаки оценки его заслуг, как количество приглашений, членство в редакциях и награды. В некоторых дисциплинах и странах определенную роль может играть субсидирование грантами. Важной составляющей общей оценки может быть и экспертная оценка, мнение коллег-ученых. Мы не должны отказываться от экспертной оценки только потому, что она иногда страдает от предвзятости; во всяком случае, у нас не меньше оснований отказаться от статистических данных, связанных с цитированием, если учесть случаи их неправильного применения. Это небольшой пример комбинации критериев, на основании которых может быть сделана оценка. Но сама система оценки, анализа во многом зависит от дисциплины. И все же сегодня «объективные» статистические данные, основанные на частоте цитирования, стали наиболее предпочтительным методом оценки. Сблaзн простого способа и несложных расчетов побеждает здравый смысл.

Авторы этой статьи – ученые-математики – хотели указать на неправильное применение количественного метода в оценке научного исследования. Особая традиция цитирования в математике, состоящая в низком количестве ссылок на журналы, статьи и авторов, делает эту науку особенно уязвимой перед статистикой. Мы полагаем, однако, что все ученые, как и широкая общественность, должны использовать правильные научные методы при оценке исследований.

Кое-кто в научном сообществе обошелся бы вообще без статистики, цитирований, но это будет означать отказ от ценного инструмента. Подсчеты, основанные на цитировании, *могут* играть роль в оценке исследования при условии, что они используются правильно, интерпретируются с осторожностью и составляют только часть анализа. Ссылки дают информацию о журналах, статьях и людях. Мы не хотим спрятать эту информацию; мы хотим пролить на нее свет.

«Все должно быть сделано как можно более просто, но не чересчур просто», – сказал однажды Альберт Эйнштейн. Этот совет одного из выдающихся ученых мира особенно уместен при оценке научных исследований.

Ранжирование журналов: импакт-фактор

Импакт-фактор создан в 1960-х гг. как способ измерять ценность журналов путем вычисления среднего числа цитирований на статью за определенный промежуток времени. Среднее число вычисляется на основе данных, собранных Thomson Scientific (бывший Институт научной информации) и публикующихся в Journal Citation Reports. Thomson Scientific ежегодно извлекает ссылки более чем из 9000 журналов, добавляя информацию о каждой статье и ее ссылках в собственную базу данных. Используя данную информацию, можно рассчитать, как часто публикация цитируется в статьях, в журналах. (Заметим, что Thomson Scientific индексирует менее половины математических журналов, охватываемых Mathematical Reviews и Zentralblatt – двумя ведущими реферативными журналами по математике.)

Величина импакт-фактора для определенного журнала в определенном году вычисляется путем расчета среднего количества ссылок на статьи за два предыдущих года в этом журнале, причем принимаются во внимание все ссылки в статьях за этот год (в рамках журналов, индексируемых Thomson Scientific). Если импакт-фактор журнала составляет 1,5 в 2007 г., то это означает, что статьи, опубликованные в нем в 2005 и 2006 гг., были 1,5 раза процитированы в статьях 2007 г., в журналах из перечня индексируемых.

Thomson Scientific использует импакт-фактор и при отборе журналов для включения в перечень индексируемых. Этот фактор нужен для того, чтобы сравнивать не только статьи, но и журналы.

Импакт-фактор может быть учтен при решении вопроса комплектации библиотек, предоставляя администратору информацию о журналах, которые уже есть в наличии, и журналах, для которых решается вопрос об их приобретении. Эти данные должны

рассматриваться наряду со сведениями о стоимости и тираже, чтобы принимать рациональные решения о покупке журналов.

Многие авторы отмечали, что нельзя судить об академической ценности журнала, используя только данные о ссылках, и авторы статьи согласны с этим мнением. В дополнение к этому общему наблюдению импакт-фактор подвергся критике и по другим причинам.

Определение импакт-фактора как среднего математического числа не совсем правильно. Поскольку многие журналы содержат редко цитируемые вспомогательные публикации, такие как письма или редакционные статьи, эти публикации не учитываются в знаменателе импакт-фактора. В то же время, пусть и нечасто, эти публикации все же цитируются и эти ссылки *учитываются* в числителе. Следовательно, импакт-фактор – не совсем среднее количество ссылок на статью. Если в журналах публикуется много таких «вспомогательных» публикаций, это отклонение может быть значительным. В некоторых областях, например в математике, это отклонение минимально.

Двухлетний период при определении импакт-фактора предназначался для того, чтобы сделать статистический показатель современным. Для некоторых областей, таких как биомедицинские науки, это целесообразно, потому что на большинство статей ссылаются вскоре после публикации. В других областях, таких как математика, большинство ссылок выходит за рамки двухлетнего периода. Анализ более 3 млн недавних ссылок в математических журналах (база данных Math. Reviews) показывает, что примерно 90 % ссылок на журнал выходит за пределы двухлетнего окна. Следовательно, импакт-фактор основывается всего лишь на 10 % ссылочной активности и игнорирует подавляющее большинство ссылок (рис. 1).

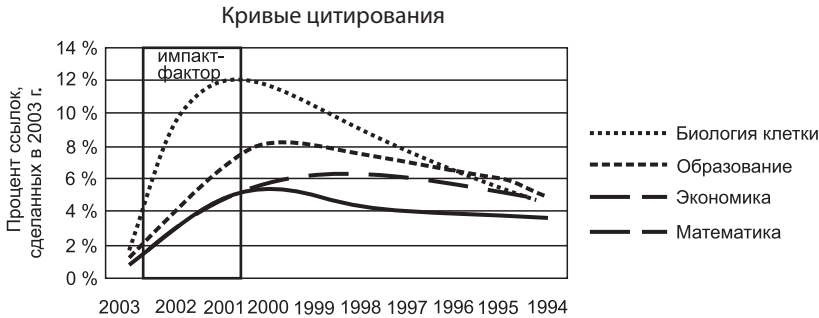


Рис. 1. Ссылочная активность в различных направлениях

График показывает возраст ссылок и охватывает четыре различные области. Ссылки на статьи, опубликованные в 2001–2002 гг. – это те, которые вносят вклад в импакт-фактор, а все остальные не имеют отношения к импакт-фактору. Данные взяты из Thomson Scientific.

Ведет ли двухлетний интервал импакт-фактора к недоразумениям? Для математических журналов ответ на этот вопрос не так однозначен. Thomson Scientific вычисляет пятилетние импакт-факторы, которые хорошо коррелируют с обычными (двухлетними) импакт-факторами. Используя базу цитирований Math. Reviews, можно вычислить «импакт-факторы» (т. е. среднее число ссылок на 1 статью) для ста наиболее цитируемых математических журналов, используя периоды в 2, 5 и 10 лет. На рис. 2 показано, что 5- и 10-летние импакт-факторы в общем соответствуют 2-летнему импакт-фактору.

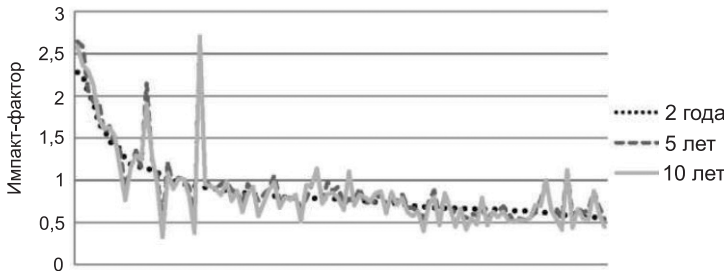


Рис. 2. Импакт-факторы для 2, 5 и 10 лет для 100 математических журналов (данные из базы цитирований Math. Reviews)

Единственным существенным выбросом на рис. 2 является журнал, который в это время не публиковал статьи. Менее значимые отклонения присущи журналам, которые публикуют небольшое количество статей каждый год, диаграмма отражает среднюю изменчивость в импакт-факторах для таких журналов. Очевидно, что изменение количества «заданных годов» при расчете импакт-фактора приводит к изменению рейтинга журналов, но они, как правило, умеренны.

Импакт-фактор значительно варьируется в зависимости от выбора дисциплины. Это различие отчасти связано со вторым наблюдением: если в некоторых дисциплинах много ссылок за пределами двухлетнего окна, то импакт-факторы для журналов ниже. Различие иногда возникает из-за того, что традиции цитирования отличаются от дисциплины к дисциплине. Из этого следует, что, используя импакт-факторы, невозможно никаким разумным способом сравнить два журнала, представляющих разные дисциплины (рис. 3).

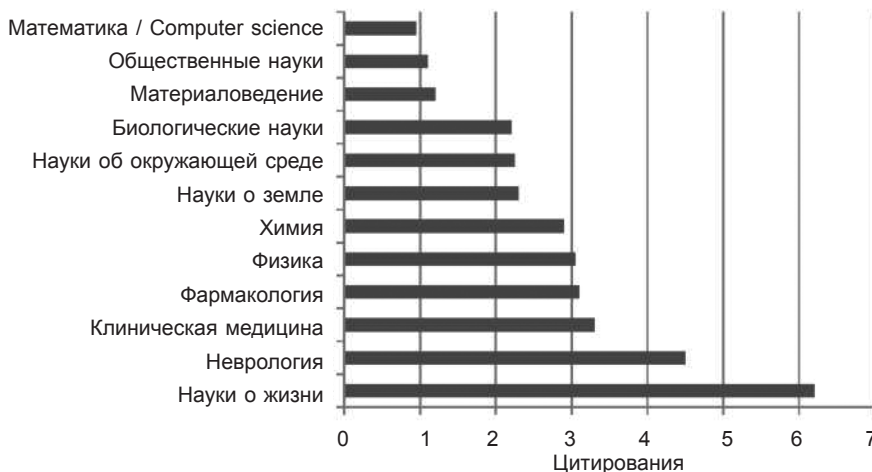


Рис. 3. Среднее количество ссылок на статью по различным дисциплинам (данные от Thomson Scientific)

Импакт-фактор может меняться от года к году, это характерно, как правило, для небольших журналов. Для журналов, публикующих меньше чем 50 статей, среднее изменение импакт-фактора с 2002 по 2003 г. составляло почти 50 %. Это полностью ожидаемо, потому что объем выборки для небольших журналов весьма мал. С другой стороны, часто сравнивают журналы за фиксированный год, игнорируя высокую годовую вариабельность данных для небольших журналов.

Журналы, которые публикуют статьи на других языках получают меньше ссылок, поскольку большая часть научного сообщества не может читать (или не читает) их. Тем самым на импакт-фактор может влиять тип журнала, язык публикаций, тираж, а не одно только качество статьи. Журналы, которые публикуют, например, обзорные статьи, часто получают больше ссылок, чем журналы, которые этого не делают, и, следовательно, у них более высокие (иногда существенно более высокие) импакт-факторы.

Наиболее существенная критика импакт-фактора в том, что не вполне ясен его смысл. При использовании импакт-фактора для сравнения двух журналов нет никакой модели для образца, в которой было бы исходно определено, что какой-то журнал «лучше». Единственным фактом считается импакт-фактор — журнал с большим импакт-фактором считается лучшим. При классическом статистическом подходе сначала определяется модель, затем формулируется гипотеза (неважно какая), а потом применяется статистическая процедура, в зависимости от результатов которой можно принять или опровергнуть гипотезу. Извлечение информации (а возможно, и построение модели) из данных — принятый способ статистического анализа, но в нашем случае не вполне ясно, как осмыслить полученную информацию. Каким образом импакт-фактор оценивает качество? Является ли он наилучшей статистической величиной для оценки качества? Что измеряет импакт-фактор?

Все приведенные выше критические замечания об импакт-факторе демонстрируют, что импакт-фактор недостаточно продуман. Например, импакт-фактор может быть использован в качестве отправной точки, чтобы разбить журналы на группы и затем с помощью других критериев составить более подробный рейтинг. Однако применение импакт-фактора для сравнительной оценки журналов требует большой осторожности. Например, с помощью импакт-фактора нельзя сравнивать журналы, относящиеся к

различным дисциплинам, необходимо учитывать тип самих журналов. Кроме того, следует уделять большое внимание годовым колебаниям импакт-фактора, особенно для небольших журналов, и иметь при этом в виду, что незначительные различия могут быть отнесены на счет случайных колебаний. Важно осознавать, что в некоторых областях знания импакт-фактор может не полностью отражать все цитирования, так как учитываются не все журналы и рассматриваемый период времени слишком короток. Статистические данные, основанные на более длительном периоде времени и большем количестве анализируемых журналов, могли бы повысить качество оценки. Наконец, цитируемость – лишь один способ оценки журналов и должна быть дополнена другой информацией (это и есть основная идея данной статьи).

Все эти соображения касаются любого упорядочивания на основе статистических данных. Необдуманная оценка журналов на основе их импакт-фактора за какой-либо год является некорректным использованием статистических данных. Следует отдать должное агентству Thomson Scientific, которое согласно этому утверждению и (в мягкой форме) предупреждает об этом всех тех, кто использует импакт-фактор: «При оценке полезности журналов Thomson Scientific учитывает не только импакт-фактор, что рекомендуется делать и другим. Импакт-фактор не может быть использован без учета других многочисленных показателей, влияющих на цитируемость, например среднего количества ссылок в одной статье. Импакт-фактор должен быть дополнен компетентной экспертной оценкой».

К сожалению, этот совет слишком часто игнорируется.

Ранжирование работ

Импакт-фактор и подобные ему статистические данные, связанные с цитируемостью, могут быть неверно истолкованы при определении рейтинга журналов. Однако существует более глубокое и более коварное заблуждение: использование импакт-фактора для сравнения отдельных работ, людей, исследовательских программ или даже целых областей знания. Эта проблема становится все актуальнее во многих странах и для многих дисциплин; она усугубляется принятыми недавно методами оценки национальных* исследований.

В определенном смысле это явление не ново. К ученым часто обращаются с просьбой оценить чей-либо список публикаций, в ответ на что регулярно слышны такие комментарии, как «она публикуется в хороших журналах» или «большая часть его работ опубликована в журналах низкого уровня». Такого рода оценки вполне разумны: качество журналов, в которых ученый обычно (или даже постоянно) публикует свои статьи, является одним из многочисленных показателей, которые могут быть использованы для общей оценки исследований данного ученого. Тем не менее импакт-фактор усилил тенденцию приписывать свойства журнала каждой статье в нем (а также каждому автору статьи).

Thomson Scientific придерживается следующей позиции: «Возможно, наиболее важным в недавно начавшемся использовании импакт-фактора является его применение для оценки научной деятельности. С помощью импакт-фактора можно составить довольно полное представление о престиже журналов, в которых публикуются ученые».

Вот некоторые примеры того, как интерпретируется эта рекомендация (по свидетельству математиков из разных частей света).

Пример 1. Недавно мой университет ввел новую классификацию журналов на основе Science Citation Index Core. Все журналы были разделены на три группы в зависимости от их импакт-фактора. В список лидеров попало 30 журналов, среди которых нет ни одного математического. Второй список содержит 667 журналов, включая 21 математический журнал. Публикация в журнале из первого списка повышает поддержку исследований со стороны университета в три раза; публикация в журналах второго списка – в два раза. Публикация в основном списке приносит 15 баллов; публикация в любом журнале перечня Thomson Scientific приносит 10 баллов. Для повышения в должности требуется определенное минимальное количество баллов.

Пример 2. В моей стране постоянные сотрудники университетов аттестуются раз в шесть лет. Несколько подряд успешно пройденных аттестаций открывают возможность для академического роста. Помимо резюме важнейшим фактором при оценке является рейтинг пяти опубликованных статей. В последние годы за каждую статью дается 3 очка, если она опубликована в журнале из первой трети списка Thomson Scientific, 2 очка, если из второй трети, и 1 очко, если статья опубликована в журна-

* Выражение «национальный» (national) общепринято в англоязычной литературе по отношению к стране (государству) в целом. – Прим. перев.

ле из последней трети (деление журналов на эти три группы происходит на основе импакт-фактора).

Пример 3. Каждый сотрудник нашего факультета оценивается при помощи формулы, включающей в себя число статей (с поправками в случае соавторства), умноженное на импакт-фактор журналов, в которых они опубликованы. На этой формуле частично основано повышение в должности и прием на работу.

В этих примерах импакт-фактор явным или неявным образом используется для сравнения отдельных статей, а также их авторов: если импакт-фактор у журнала *A* больше, чем у журнала *B*, то, безусловно, статья из *A* должна быть лучше статьи из *B*, а автор статьи из *A* лучше автора статьи из *B*. В некоторых случаях это рассуждение распространяется и на оценку факультетов или даже целых областей знания.

Давно известно, что распределение числа цитирований по статьям в одном журнале крайне неравномерно и близко к так называемому степенному закону. Последствия этого продемонстрируем на примере.

Приведем распределение цитируемости по статьям в журнале *Proceedings of the American Mathematical Society* (*Proceedings of the AMS* или *Proceedings*) за период 2000–2004 гг. (рис. 4). *Proceedings* публикует короткие статьи, менее десяти страниц каждая. За этот период журналом опубликована 2381 работа (около 15 000 страниц). На основе данных *Math. Reviews* о цитировании за 2005 г. среднее количество цитирований одной статьи (импакт-фактор) 0,434.

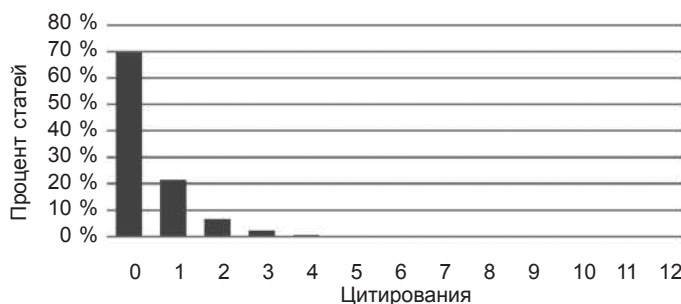


Рис. 4. Распределение цитируемости по статьям в журнале *Proceedings of the AMS* за 2000–2004 гг.

Журнал *Transactions of the American Mathematical Society* (*Transactions of the AMS* или *Transactions*) публикует более длинные статьи, которые, как правило, значительнее как по объему, так и по содержанию, чем статьи из *Proceedings*. За тот же период времени *Transactions* опубликовал 1165 работ (более 25 000 страниц) с числом цитирований от 0 до 12. Среднее количество цитирований одной статьи 0,846, что примерно в два раза больше, чем для *Proceedings*.

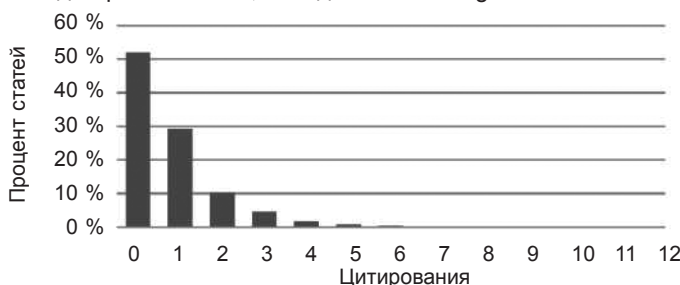


Рис. 5. Распределение цитируемости по статьям в журнале *Transactions of the AMS* за 2000–2004 гг.

Теперь предположим, что из двух математиков один опубликовал статью в *Proceedings*, а другой — в *Transactions*. При упомянутых выше подходах второй математик должен быть оценен выше первого, поскольку статья второго математика опубликована в журнале с большим импакт-фактором (причем в два раза!). Но насколько правомерно такое мнение? Действительно ли статьи из *Transactions of the AMS* в два раза лучше статей из *Proceedings*?

Когда мы утверждаем, что статьи из *Transactions* лучше (в смысле цитируемости) статьи из *Proceedings*, мы должны сравнивать не средние статьи в этих журналах, а соответствующие вероятности: какова вероятность того, что мы не правы? То есть какова вероятность того, что у случайно выбранной статьи из *Proceedings* по крайней мере столько же цитирований, сколько у случайно выбранной статьи из *Transactions*?

Элементарное вычисление дает величину 62 %. Это означает, что мы неправы в 62 % случаев и случайно выбранная статья из *Proceedings* будет по крайней мере не хуже случайно выбранной статьи из *Transactions*, несмотря на то что импакт-фактор у *Proceedings* равен всего лишь половине импакт-фактора у *Transactions*! Таким образом, мы чаще ошибаемся, чем оказываемся правы. Большинству людей это кажется удивительным, хотя это всего лишь следствие асимметричного распределения цитируемости по статьям и слишком короткого интервала времени, используемого для вычисления импакт-фактора (причина высокого процента нецитируемых работ). Это показывает важность точного статистического анализа по сравнению с интуитивным представлением.

Ситуация эта характерна для всех журналов; два выбранных нами журнала не являются какими-то особенными. (Например, *Journal of the AMS* за тот же период имеет импакт-фактор 2,63, что в шесть раз больше, чем у *Proceedings*. Однако в 32 % случаев выбранная наугад статья из *Proceedings* в смысле цитирования ничуть не хуже статьи из *Journal of the AMS*.)

Таким образом, хотя и неверно было бы утверждать, что импакт-фактор совсем ничего не говорит об отдельных статьях в журнале, он дает весьма расплывчатую информацию и может ввести в глубокое заблуждение.

Коль скоро мы осознаем бессмысленность импакт-фактора для оценки количества цитирований отдельной статьи, не имеет смысла применять его и для оценки авторов статей, исследовательских программ, которыми они занимаются, и тем более областей знаний, в которых они работают. Импакт-фактор, как и вообще средние показатели, не слишком пригодная величина для содержательного сравнения.

Конечно, ранжирование людей не то же, что ранжирование их статей. Однако если требуется оценить чьи-либо статьи, принимая цитируемость как меру качества, то начать следует с количества цитирований каждой статьи. Импакт-фактор журнала, в котором опубликована статья, неадекватная замена этого показателя.

Ранжирование ученых

Хотя импакт-фактор – наиболее известный метод статистики, основанный на цитировании, недавно появились и другие показатели, которые сейчас активно популяризируются.

Вот три из них.

h-индекс: у ученого есть n статей, на каждую из которых существует по крайней мере n ссылок. Это самая популярная из упомянутых выше статистических величин. Ее ввел Дж. Е. Хирш в попытке оценить количественно научную продукцию ученого, приняв за основу «хвост» кривой цитируемости его статей.

m-индекс: отношение h -индекса ученого к числу лет, прошедших после его первой публикации. Этот показатель предложен тем же Хиршем. Учет m -индекса позволяет компенсировать молодым ученым недостаток времени на публикацию многих работ и набрать больше цитирований.

g-индекс: определяется как наибольшее p , для которого p наиболее цитируемых работ автора в общей сложности цитируются по крайней мере p^2 раз. Этот показатель предложил Лео Эгг в 2006 г. Поскольку h -индекс не принимает во внимание тот факт, что некоторые из p наиболее цитируемых статей могут быть очень высоко цитируемыми, g -индекс призван исправить этот недостаток.

Существуют и другие индексы, и их много. В их числе модификации названных выше, в которых учитывается возраст работ или количество авторов.

В своей работе Хирш предлагает h -индекс как легко вычисляемую величину, которая дает оценку важности, значимости и масштаба влияния совокупного вклада ученого в науку. Далее он добавляет, что «этот показатель может служить полезным средством для сравнения различных конкурирующих лиц, когда критерием оценки являются научные достижения».

Ни одно из этих утверждений не подтверждается убедительными доказательствами. В поддержку своего утверждения о важности и значимости h -индекса, совокупного вклада ученого в науку, Хирш приводит данные анализа h -индекса для выборки из лауреатов Нобелевской премии (и членов Национальной академии). Он показывает, что члены этих групп, как правило, имеют высокий h -индекс. И все же без дополнительной информации мы вряд ли способны сделать заключение о том, что некто станет Нобелевским лауреатом или членом Национальной академии, если известно только, что у него высокий h -индекс.

В своей статье Хирш также утверждает, что можно использовать h -индекс для сравнения двух ученых: два человека, имеющих одинаковую величину h , сопоставимы с точки зрения их научного вклада, даже если количество их статей или цитиро-

ваний очень разные. И наоборот, из двух людей (одного и того же научного возраста) с одинаковым количеством работ или цитирований и сильно различающимися h -значениями тот, у которого h больше, вероятно, ученый более высокого уровня.

Эти утверждения не в ладах со здравым смыслом. (Представьте себе двух ученых, у каждого из которых есть по 10 работ и на каждую из них есть по 10 ссылок, но при этом у одного есть еще 90 работ, на каждую из которых ссылаются по 9 раз. Или, например, у одного 10 работ, и на каждую из них ссылаются по 10 раз, а у другого – 10 работ, и на каждую из них ссылаются по 100 раз. Неужели кто-то и вправду будет считать этих людей равными?)

Хирш превозносит достоинства h -индекса, утверждая, что h -индекс предпочтительнее других численных критериев, которые обычно используются для оценки научной деятельности исследователей, но не определяет смысл слова «предпочтительнее», как и не объясняет, почему *хотят* найти именно «критерий с одним показателем».

Хотя данный подход и подвергался критике, но ему недоставало серьезного анализа. Большая часть анализа демонстрировала «сходящуюся валидность», т. е. то, что h -индекс хорошо коррелирует с другими методами оценки публикаций, такими как количество опубликованных работ или общее число цитирований. Такая корреляция не очень впечатляет, так как все эти величины – функции одной и той же основной сущности – самих публикаций. В одной интересной статье об h -индексе проводится более тщательный анализ и показывается, что h -индекс (точнее говоря, m -индекс) «хуже», чем простое среднее ссылок на работы. Отметим, что даже здесь авторы не в состоянии адекватно определить, что значит «хуже». При применении классической статистической парадигмы h -индекс оказывается менее надежным, чем другие показатели.

Были разработаны разные варианты h -индекса, приспособленные для сравнения качества исследователей не только в рамках одной дисциплины, но и для сравнения качества исследований в различных областях. Некоторые полагают, что h -индекс может быть использован для сравнения различных институтов и их подразделений. В основном все это – наивные попытки описать весь сложный процесс цитирования с помощью одного показателя. Действительно, основное преимущество этих новых показателей над обычными гистограммами цитирований в том, что индекс избавляет нас от излишних деталей, что делает любых двух исследователей сравнимыми между собой. Однако даже на простых примерах видно, что отброшенная информация необходима, чтобы понимать сущность исследования. Именно понимание должно быть целью оценки исследований, а не только обеспечение того, чтобы любые два человека могли быть сравнимы.

В некоторых случаях h -индекс или его варианты применяются государственными органами, оценивающими исследования. Это не что иное, как использование данных не по назначению. К сожалению, слишком велик соблазн определить рейтинг каждого ученого, оперируя всего одним показателем; может статься, что этим соблазнится и широкая общественность, которая и в более простых случаях неспособна понять, что следует из тех или иных статистических величин. Те, кто продвигает мысль о продвижении метода статистики цитирований в качестве основной меры качества исследований, не дают ответа на главный вопрос: что означает цитирование? Они собирают большое количество данных о цитировании, обрабатывают их, получают из них статистические показатели, а затем утверждают, что оценки на этой основе «объективны». Однако мы можем лишь *интерпретировать* статистические показатели, а интерпретация основана на понимании *смысла* цитирования, она субъективна.

Как ни странно, в литературе в поддержку такого подхода трудно найти четкие аргументы «за».

- Идея, стоящая за индексацией цитирований, принципиально проста. Если признать, что ценность информации определяется теми, кто ее использует, что может быть лучше для оценки качества работы, чем измерение ее влияния на сообщество.

- Хотя измерить одним показателем качество отдельных ученых затруднительно, общее мнение в том, что лучше публиковать больше, чем меньше, и что количество ссылок на работу (с учетом общих традиций цитирования в данной области) есть адекватная мера ее качества.

- Частота цитирования отражает качество журнала и пользу, которую из него извлекают.

- Когда врач или специалист в области биомедицины цитирует журнальную статью, это указывает, что цитируемый журнал каким-то образом повлиял на него.

- Цитаты – признание интеллектуального долга.

Важные для нас термины – «качество», «значение», «влияние» и «интеллектуальный долг». Термин «импакт» стал родовым понятием, придающим цитированию вес и смысл. Термин впервые возник в короткой статье, которую написал в 1955 г. Е. Гарфилд, продвигая идею индекса цитирования. По его словам, в случае весьма значимых работ индекс цитирования имеет количественное значение, поскольку может помочь историку в измерении степени влияния работы, т. е. ее импакт-фактора.

Довольно ясно, что здесь, как во всех других случаях, термин «импакт-фактор», или «степень влияния», предполагает, что цитирующая работа «построена» на цитируемой, т. е. цитирование – тот механизм, посредством которого исследования движутся вперед.

Обширная литература на эту тему демонстрирует, что процесс цитирования более сложен, чем можно предположить, опираясь на эти расплывчатые заявления. Например, в работе 1983 г. об оценке исследований Мартин и Ирвайн пишут, что в основе всех этих проблем с использованием цитирования в качестве меры качества лежит наше непонимание причин того, почему авторы ссылаются на одни работы, а не на другие. По их мнению, простой анализ цитирований предполагает опору на весьма рациональную модель цитирования, согласно которой цитирование происходит по преимуществу с учетом научной признанности предыдущей работы, имеющей высокое качество или важность, и у данной работы существует равная вероятность быть процитированной разными авторами.

В своей работе 1988 г. о смысле цитирования Коззенс утверждает, что цитирование – результат одной из двух линий поведения при оформлении научных публикаций – «признательной» и «риторической».

Цитирования первого типа несут в себе признание того, что цитирующая работа имеет «интеллектуальный долг» по отношению к цитируемой. Цитирования второго типа просто ссылка на предшествующую работу, объясняющую какой-то результат, быть может, даже вовсе не принадлежащий цитируемому автору. Такие «риторические» цитирования – средство ведения научного обсуждения, а не признания интеллектуального долга. Конечно, в некоторых случаях цитирование может представлять собой оба типа.

По наблюдениям Коззенс, большинство цитирований имеют «риторический» характер.

Это подтверждается опытом большинства практикующих математиков. (Например, в базе данных *Math. Reviews* из более чем 3 млн ссылок почти 30 % приходится на книги, а не на исследовательские статьи в журналах.) Почему это важно? Потому что в отличие от «признательных» ссылок, которые относятся к фундаментальным работам, выбор, какие работы цитировать «риторически», зависит от многих факторов: известности цитируемого автора, отношения между цитирующим и цитируемым, доступности соответствующего журнала (все ли журналы в открытом доступе более цитируемы?), желания для удобства сослаться на несколько результатов одной работы и т. д. Немногие из этих факторов имеют непосредственное отношение к «качеству» цитируемой работы.

Даже когда цитирование носит «признательный» характер, оно может быть обусловлено различными мотивами, в том числе степенью известности, отрицательной репутацией, рабочим характером информации, убедительностью, положительной репутацией, желанием предупредить читателя или общественным мнением. В большинстве случаев цитирование обусловлено более чем одним из факторов. Некоторые заметные результаты могут пострадать из-за эффекта «стирания», если сразу же включаются в другие работы, которые и служат основой для ссылок.

Некоторые ссылки не просто выражают «признательность» за выдающиеся исследования, а служат предупреждением об ошибочных результатах. Можно привести много примеров таких «предупреждающих» цитирований.

Социология цитирования – сложный предмет, который выходит за рамки настоящего доклада. Однако даже наше беглое обсуждение показывает, что цитирование имеет статистические погрешности и не так «объективно», как утверждают их сторонники.

Некоторые могут возразить, что поиск смысла цитирований не играет особой роли, поскольку статистические показатели, основанные на цитировании, хорошо коррелируют с некоторыми другими мерами качества исследований (скажем, экспертной оценкой).

Например, упоминавшееся ранее сообщение в *Evidence* утверждает, что основанные на цитировании статистические показатели могут (и должны) заменить другие формы оценки.

Судя по всему, из этого делается заключение, что основанные на цитировании статистические показатели, независимо от их точного смысла, должны заменить другие методы оценки, потому что их хорошо учитывать. Нелепость такого предложения всякому очевидна.

Разумное использование статистики

Ревностные попытки чрезмерно довериться объективным метрикам (статистическим показателям) для оценки исследований не новы. Этот метод оценки красноречиво описан в вышедшей в 2001 г. популярной книге *Damned lies and statistics* («Проклятая ложь и статистика») социолога Дж. Беста: «Есть культуры, представители которых верят, что некоторые объекты имеют магическую силу; антропологи называют эти объекты фетишами. В нашем обществе своего рода фетиш – статистика. Мы склонны придавать статистическим показателям магический смысл, как будто они что-то большее, чем просто числа. Мы относимся к ним как к образцу истины, как если бы они перерабатывали сложность и запутанность реальности в простоту и ясность. Мы используем их, чтобы свести сложные социальные проблемы к более понятным оценкам, процентам и отношениям».

Статистика дает направление решению проблем, она показывает нам, о чем и насколько сильно мы должны беспокоиться. В некотором смысле социальная проблема становится и статистической, и поскольку мы относимся к статистике как к истине в последней инстанции, именно статистика становится фетишем, имеющим магический контроль над тем, как мы рассматриваем социальные проблемы. Мы считаем, что статистика – это факты, которые мы открыли, а не просто выдуманные нами самими числа».

Эту мистическую веру в волшебную силу статистики цитирований можно найти в документации для программ, как национальных, так и ведомственных, оценки качества исследовательской работы, в работах, использующих *h*-индекс и его модификации.

Это отношение проявляется в современных попытках улучшить импакт-фактор при анализе цитат путем использования более сложных математических алгоритмов, в том числе алгоритмов классификации страниц. Сторонники этого подхода заявляют о его эффективности, что не оправдывается анализом и плохо поддается оценке. Более сложные расчеты, предположения (зачастую скрытые), стоящие за ними, многим людям не так просто разглядеть. Предполагается, что мы должны относиться к цифрам и рейтингам с благоговением – как к истине, а не как к тому, что создано нами.

Научные исследования – не первый вид публично финансируемой деятельности, подвергающейся тщательному рассмотрению: за последние десятилетия люди пытались произвести количественные оценки производительности всего, начиная с системой образования (школы) и до здравоохранения (больницы и даже отдельных хирургов).

В некоторых случаях к делу подключались специалисты по статистике, чтобы проконсультировать всех и каждого по поводу разумных подсчетов и правильного использования статистических данных. Если в практической медицине принято консультироваться с врачами, то в области статистики, безусловно, следует советоваться со статистиками (и прислушиваться к их советам). Два прекрасных примера можно найти в работах. Хотя они имеют дело с оценкой производительности, а не с научными исследованиями (мониторинг производительности государственного сектора и здравоохранения/образования во втором), каждый дает представление о разумном использовании статистики в оценке научных исследований.

Работа Гольдштейна и Шпигельхайтера имеет дело с использованием *League Tables* (рейтингов) на основе простейших числовых показателей (например, достижений студентов или медицинских результатов), и в особенности это относится к оценке научных исследований с помощью рейтинга журналов, статей или авторов посредством статистики цитирований.

Данные

Каков бы ни был мыслимый объем собранных статистических данных, он не сможет преодолеть ключевых сомнений относительно пригодности или полноты этих данных.

Это важное замечание для оценок производительности, основанных на цитировании.

Импакт-фактор, например, основан только на тех данных, которые получены по журналам из перечня Thomson Scientific. (Заметим, что импакт-фактор сам по себе является основным критерием отбора журналов в этот перечень.) Некоторые ставят под сомнение целостность таких данных. Другие указывают на то, что иные наборы данных могут быть более полными. Некоторые группы продвинули идею использования Google Scholar для внедрения статистики, основанной на цитировании, например *h*-индекса, но данные, содержащиеся в Google Scholar, часто неточны (так как, например, имена авторов автоматически извлекаются из публикаций в интернете). Индекс цитирования отдельных ученых иногда трудно получить, поскольку в некоторых ситуациях бывает непросто однозначно определить авторство, и это может стать огромным препятствием для сбора точных данных по цитатам. Зачастую не прини-

маются во внимание некоторые данные, которые относятся к анализу цитирования. А из статистики, основанной на ошибочных данных, делаются ошибочные выводы.

Мы обратим особое внимание на выбор подходящей статистической модели, на досадную *неопределенность* в представлении всех результатов, на методы коррекции результатов с учетом возмущающих факторов и, наконец, на то, до какой степени мы можем доверять имеющимся *рейтингам*.

Как мы уже писали ранее, в большинстве случаев, в которых метод подсчета цитирования используется для оценки работ, людей и программ, конкретной модели заранее не определяется. Сами данные подводят к весьма расплывчатой идее. Образуется порочный круг, когда объекты оцениваются выше, потому что у них (в базе данных) более высокий рейтинг. Неопределенности *любого* из этих рейтингов уделяется недостаточное внимание, и мало анализируется, как эта неопределенность (например, ежегодные изменения импакт-фактора) будет влиять на рейтинги. Наконец, второстепенные факторы (например, конкретная дисциплина, тип статей, которые публикует журнал, является ли конкретный ученый экспериментатором или теоретиком) часто игнорируются в таких рейтингах, особенно если они оценивают результаты в национальном масштабе.

Интерпретация и влияние

Тема, которая обсуждается в данной работе, вызывает большой общественный интерес, и, очевидно, это та область, где особое внимание к недостаткам жизненно важно, но чаще всего игнорируется. Являются ли скорректированные результаты справедливой мерой «качества» какого-либо учреждения – это вопрос; с другой стороны, аналитики должны также знать о потенциальном влиянии результатов на изменения в поведении учреждений и отдельных лиц, желающих улучшить свой последующий «рейтинг».

Оценка научных исследований также вызывает большой общественный интерес. Результаты оценки отдельного ученого могут оказать серьезное и долгосрочное влияние на его карьеру; оценка отдела может изменить шансы на успех в далеком будущем, а что касается дисциплин, то набор оценок может разделить их на процветающие и «застойные». Поскольку дело столь важно, надо, конечно, отдавать себе отчет в качестве метода подсчета.

В какой степени цитирование измеряет качество исследований? Кажется, что количество цитирований коррелирует с качеством, и все, конечно, понимают, что высококачественные статьи много цитируются. Но, как говорилось выше, некоторые статьи, особенно в определенных дисциплинах, много цитируются по причинам, не потому что их качество высоко. Необходимо лучше понять точную природу рейтингов на основе метода подсчета цитирований. Кроме того, если статистика цитирований играет центральную роль в оценке научных исследований, ясно, что авторы, редакторы и даже издатели найдут способы управлять системой в своих интересах. Долгосрочные последствия этого неясны и не изучены.

Статью Гольдштейна и Шпигельхальтера очень ценно прочитать сегодня, потому что из нее вытекает, что чрезмерная зависимость от простой статистики в оценке научных исследований не решает проблему. Правительства, учреждения и частные лица боролись с подобным в прошлом и нашли способы лучше понять статистические инструменты и дополнить их другими средствами оценки. Гольдштейн и Шпигельхальтер заканчивают свою работу в позитивном ключе: «Наконец, хотя мы в целом настроены критично ко многим нынешним попыткам составить мнения об учреждениях, мы не хотим создать впечатление, что считаем, что все такие методики сравнения не имеют смысла. Нам кажется, что сравнение учреждений и попытка понять, почему они отличаются, – чрезвычайно важное направление деятельности и лучше проводить его в духе сотрудничества, а не конфронтации».

Пожалуй, это единственный надежный способ получения объективной информации, которая может привести к пониманию и к совершенствованию процедуры. *Реальной проблемой с упрощенными процедурами, которые мы подвергли критике, является то, что они отвлекают от этой достойной цели как внимание, так и ресурсы*».

Было бы трудно найти лучшие слова, чтобы сформулировать цели, которые должны ставить перед собой все те, кто причастен к организации оценки исследований.

Примечания

1) В Reader's Digest от октября 1977 г. эта цитата была приписана Эйнштейну. Похоже, что она была извлечена из этой цитаты: «Едва ли можно отрицать, что высшая цель всех теорий – сделать основные элементы как можно более простыми и немногочисленными, не поступаясь при этом полным анализом проблемы». Из

Спенсеровской лекции «О методе теоретической физики», прочитанной Эйнштейном в Оксфорде (10 июня 1933 г.); опубликовано также в *Philosophy of Science* 1 (1934), №. 2, 163—169.

2) Хотя в этом разделе мы сосредоточились на импакт-факторе Thomson Scientific, отметим, что Thomson продвигает и другие методы учета. Добавим, что аналогичные методики, основанные на подсчете среднего числа цитирований в журналах, могут быть получены из других баз данных, включая Scopus, Spires, Google Scholar и (по математике) Math. Reviews. Последняя включает цитаты из более чем 400 математических журналов с 2000 г. по настоящее время. База содержит более 3 млн ссылок.

3) Thomson Scientific указывает (март 2008 г.), что рассматривает журналы в следующих категориях:

- математика (217);
- прикладная математика (177);
- междисциплинарная математика (76);
- математическая физика (444);
- вероятность и статистика (96).

Категории перекрываются, а общее число журналов составляет около 400.

Math. Reviews каждый год анализирует более 1200 журналов, из которых более 800 журналов составляют «ядро» (в том смысле, что каждая статья в журнале включена в Math. Reviews). Zentralblatt охватывает примерно столько же математических журналов.

4) База данных цитирований в Math. Reviews включает (по состоянию на март 2008 г.) более 3 млн ссылок приблизительно из 400 журналов, изданных с 2000 г. по настоящее время. Ссылки согласованы с разделами в базе данных Math. Reviews и охватывают период в несколько десятков лет. В отличие от Science Citation Index сюда включено цитирование и из книг, и из журналов. Любопытно, что примерно 50 % цитирований относится к публикациям предыдущего десятилетия, 25 % ссылок приходится на статьи, появившиеся в десятилетие до него, 12,5 % – в предшествующее десятилетие и т. д. Для каждой дисциплины характерна, конечно, своя картина.

5) Высокая неравномерность распределения цитируемости по статьям в сочетании с коротким интервалом времени (рассматриваются ссылки на статьи, опубликованные за последние пять лет, в статьях, опубликованных за последний год) означает, что много статей имеют либо мало ссылок, либо не имеют их вообще. Интуитивно понятно, что это приводит к равному количеству цитирований у двух произвольно выбранных статей.

Тот факт, что многие статьи не имеют цитирований, также следствие большого временного интервала, характерного для цитирования в математических статьях: часто перед появлением первого цитирования проходит много лет. Если рассматривать более длительный период времени как для источников цитирований, так и для самих статей, то количество цитирований заметно возрастет и станет легче различать журналы по этому показателю. Таким образом, появляется механизм, позволяющий сравнивать журналы в терминах соответствующих распределений, что, безусловно, более продумано, чем использование импакт-фактора. И тем не менее при этом опять-таки учитывается только цитируемость и ничего более.

6) Чтобы показать, как много информации теряется при применении только *h*-индекса, приведем пример одного уважаемого математика, находящегося в середине своей карьеры, который опубликовал 84 научные работы (рис. 6).

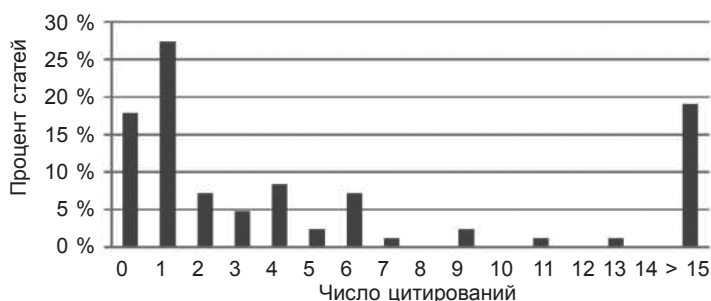


Рис. 6. Учет цитирований ученого (84 статьи) с применением только *h*-индекса

На рис. 6 видно, что чуть менее 20 % публикаций имеют 15 или более цитирований. Распределение действительных подсчетов цитирований для этих 15 работ представлено на рис. 7.

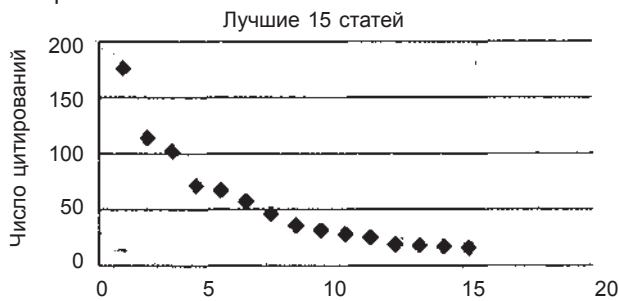


Рис. 7. Распределение действительных подсчетов цитирований для 15 лучших статей

В анализе Хирша, однако, вся эта информация выброшена. Он лишь констатирует, что *h*-индекс равен 15, т. е. что 15 лучших работ имеют 15 или более цитирований.

Алгоритм из работы [Bergstrom 2007] использует алгоритм страничного ранжирования, который придает каждой цитате вес, а затем вычисляет импакт-фактор. Сильной стороной алгоритмов страничного ранжирования является то, что они принимают во внимание «ценность» цитат.

С другой стороны, их сложность может таить опасность. В этом случае все «самоцитирования» отбрасываются, т. е. отбрасываются все цитаты из статей в данном журнале на статьи, опубликованные в том же журнале за предыдущие пять лет. По сути, это нельзя назвать «самоцитированием», но рассмотрение некоторых данных базы цитат Math. Reviews позволяет предположить, что это приводит к выбрасыванию трети всех цитирований.

Алгоритм из работы [Springer et al. 2008] интересен отчасти потому, что он пытается рассмотреть различные временные рамки цитирований, а также сравнить случайно выбранные статьи в одном журнале со статьями в другом. В этом случае сложность алгоритмов затрудняет для большинства людей оценки результатов.

Поступила в редакцию 20.04.2014.

УДК 37.0+316.7+13

Л. Г. ТИТАРЕНКО,
ДОКТОР СОЦИОЛОГИЧЕСКИХ НАУК, ПРОФЕССОР (МИНСК)

НОВЫЕ И СТАРЫЕ ПРОБЛЕМЫ КАЧЕСТВА ОБРАЗОВАНИЯ В БЕЛАРУСИ

Рассматриваются проблемы повышения качества высшего образования в Республике Беларусь. Дается анализ компонентов образования, от которых зависит его качество. Показано, какова роль прежних и нынешней реформы высшего образования в повышении его конкурентоспособности и приведении белорусского высшего образования в соответствие с задачами развития общества и личности. Предложены пути улучшения ситуации с преподаванием социально-гуманитарных дисциплин.

Ключевые слова: высшее образование; реформы; качество обучения; потребности студентов; философия современного образования; инновации в образовании.

The article is devoted to the problems of increasing the quality of higher education in the Republic of Belarus. The components of education making up its quality is given analysis to. The author shows the role of the previous and current reforms of higher education in increasing its competitiveness and bringing the higher education system to conformity with the current aims of societal development and personality. Some ways to improve the situation with social-humanitarian disciplines are suggested.

Key words: higher education; reforms; quality of education; students' needs; philosophy of modern education; innovations in education.

Современное общество находится сегодня в чрезвычайно сложных условиях развития. Эта сложность имеет место во всем мире ввиду его глобализации и роста взаимосвязи между происходящими в нем процессами, странами, регионами и т. п.